

#2

0590

10/09 OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/841,805

DATE: 06/08/2001
 TIME: 11:46:26

Input Set : C:\PAOLA\09841805.txt
 Output Set: N:\CRF3\06082001\I841805.raw

SEQUENCE LISTING

ENTERED

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Lal, Preeti
 6 Shah, Purvi
 7 Corley, Neil C.
 9 (ii) TITLE OF INVENTION: HUMAN MEMBRANE RECYCLING
 10 PROTEINS

12 (iii) NUMBER OF SEQUENCES: 5

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 16 (B) STREET: 3174 Porter Dr.
 17 (C) CITY: Palo Alto
 18 (D) STATE: CA
 19 (E) COUNTRY: USA
 20 (F) ZIP: 94304

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette
 24 (B) COMPUTER: IBM Compatible
 25 (C) OPERATING SYSTEM: DOS
 26 (D) SOFTWARE: FastSEQ for Windows Version 2.0

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/841,805

C--> 30 (B) FILING DATE: 24-Apr-2001

31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 09/360,125
 35 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Billings, Lucy J.
 39 (B) REGISTRATION NUMBER: 36,749
 40 (C) REFERENCE/DOCKET NUMBER: PF-0456 US

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 650-855-0555
 44 (B) TELEFAX: 650-845-4166

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 347 amino acids
 51 (B) TYPE: amino acid
 52 (C) STRANDEDNESS: single
 53 (D) TOPOLOGY: linear

55 (vii) IMMEDIATE SOURCE:

56 (A) LIBRARY: TONGTUT01
 57 (B) CLONE: 980615

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 Met Ala Gln Ser Arg Asp Gly Gly Asn Pro Phe Ala Glu Pro Ser Glu
 62 1 5 10 15
 63 Leu Asp Asn Pro Phe Gln Asp Pro Ala Val Ile Gln His Arg Pro Ser

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,805

DATE: 06/08/2001

TIME: 11:46:26

Input Set : C:\PAOLA\09841805.txt

Output Set: N:\CRF3\06082001\I841805.raw

```

64          20          25          30
65 Arg Gln Tyr Ala Thr Leu Asp Val Tyr Asn Pro Phe Glu Thr Arg Glu
66          35          40          45
67 Pro Pro Pro Ala Tyr Glu Pro Pro Ala Pro Ala Pro Leu Pro Pro Pro
68          50          55          60
69 Ser Ala Pro Ser Leu Gln Pro Ser Arg Lys Leu Ser Pro Thr Glu Pro
70 65          70          75          80
71 Lys Asn Tyr Gly Ser Tyr Ser Thr Gln Ala Ser Ala Ala Ala Ala Thr
72          85          90          95
73 Ala Glu Leu Leu Lys Lys Gln Glu Glu Leu Asn Arg Lys Ala Glu Glu
74          100         105         110
75 Leu Asp Arg Arg Glu Arg Glu Leu Gln His Ala Ala Leu Gly Gly Thr
76          115         120         125
77 Ala Thr Arg Gln Asn Asn Trp Pro Pro Leu Pro Ser Phe Cys Pro Val
78          130         135         140
79 Gln Pro Cys Phe Phe Gln Asp Ile Ser Met Glu Ile Pro Gln Glu Phe
80 145          150          155          160
81 Gln Lys Thr Val Ser Thr Met Tyr Tyr Leu Trp Met Cys Ser Thr Leu
82          165          170          175
83 Ala Leu Leu Leu Asn Phe Leu Ala Cys Leu Ala Ser Phe Cys Val Glu
84          180          185          190
85 Thr Asn Asn Gly Ala Gly Phe Gly Leu Ser Ile Leu Trp Val Leu Leu
86          195          200          205
87 Phe Thr Pro Cys Ser Phe Val Cys Trp Tyr Arg Pro Met Tyr Lys Ala
88          210          215          220
89 Phe Arg Ser Asp Ser Ser Phe Asn Phe Phe Val Phe Phe Phe Ile Phe
90 225          230          235          240
91 Phe Val Gln Asp Val Leu Phe Val Leu Gln Ala Ile Gly Ile Pro Gly
92          245          250          255
93 Trp Gly Phe Ser Gly Trp Ile Ser Ala Leu Val Val Pro Lys Gly Asn
94          260          265          270
95 Thr Ala Val Ser Val Leu Met Leu Leu Val Ala Leu Leu Phe Thr Gly
96          275          280          285
97 Ile Ala Val Leu Gly Ile Val Met Leu Lys Arg Ile His Ser Leu Tyr
98          290          295          300
99 Arg Arg Thr Gly Ala Ser Phe Gln Lys Ala Gln Gln Glu Phe Ala Ala
100 305          310          315          320
101 Gly Val Phe Ser Asn Pro Ala Val Arg Thr Ala Ala Ala Asn Ala Ala
102          325          330          335
103 Ala Gly Ala Ala Glu Asn Ala Phe Arg Ala Pro
104          340          345
106 (2) INFORMATION FOR SEQ ID NO: 2:
108 (i) SEQUENCE CHARACTERISTICS:
109 (A) LENGTH: 1521 base pairs
110 (B) TYPE: nucleic acid
111 (C) STRANDEDNESS: single
112 (D) TOPOLOGY: linear
114 (vii) IMMEDIATE SOURCE:
115 (A) LIBRARY: TONGTUT01

```

RAW SEQUENCE LISTING

DATE: 06/08/2001

PATENT APPLICATION: US/09/841,805

TIME: 11:46:26

Input Set : C:\PAOLA\09841805.txt

Output Set: N:\CRF3\06082001\I841805.raw

```

116          (B) CLONE: 980615
117
118          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
120  NGACGCAGGC  GCAACCCACG  GCTGCTGCGG  GGATCCTTGT  GGCCCTTCCG  GTCGATGGAA      60
121  CCAATCCGTG  CACAGAGAAG  CGGGGCGAAC  TGAGGCGAGT  GAAGTGGAAT  CTGAGGGGCTA    120
122  CCGCTACCGC  CACTGCTGCG  GCAGGGGCGT  GGAGGGCAGA  GGGCCGCGGA  GGCCGCAGTT     180
123  GCAAACATGG  CTCAGAGCAG  AGACGGCGGA  AACCCGTTCT  CCGAGCCGAG  CGAGCTTGAC     240
124  AACCCTTTTC  AGGACCCAGC  TGTGATCCAG  CACCGACCCA  GCCGGCAGTA  TGCCAACGCTT    300
125  GAGCTCTACA  ACCCTTTTGA  GACCCGGGAG  CCACCACCAG  CCTATGAGCC  TCCAGCCCCT     360
126  GCCCCATTGC  CTCCACCCCT  AGCTCCCTCC  TTGCAGCCCT  CGAGAAAAGCT  CAGCCCCACA     420
127  GAACCTAAGA  ACTATGGGCT  ATACAGCACT  CAGGCCTCAG  CTGCAGCAGC  CACAGCTGAG     480
128  CTGCTGAAGA  AACAGGAGGA  GCTCAACCGG  AAGGCAGAGG  AGTTGGACCG  AAGGGAGCGA     540
129  GAGCTGCAGC  ATGCTGCCCT  GGGGGGCACA  GCTACTCGAC  AGAACAATTG  GCCCCCTCTA     600
130  CCTTCTTTTT  GTCCAGTTCA  GCCCTGCTTT  TTCCAGGACA  TCTCCATGGA  GATCCCCCAA     660
131  GAATTTTCTA  GACTGTATC  CACCATGTAC  TACCTCTGGA  TGTGCAGCAC  GCTGGCTCTT     720
132  CTCCTGAACT  TCCTCGCCTG  CCTGGCCAGC  TTCTGTGTGG  AAACCAACAA  TGGCGCAGGC     780
133  TTTGGGCTTT  CTATCCTCTG  GGTCTCCTT  TTTACTCCCT  GCTCCTTTGT  CTGCTGGTAC     840
134  CGCCCCATGT  ATAAGGCTTT  CCGGAGTGAC  AGTTCATTCA  ATTTCTTCGT  TTTCTTCTTC     900
135  ATTTTCTTCG  TCCAGGATGT  GCTCTTTGTC  CTCCAGGCCA  TTGGTATCCC  AGGTTGGGGA     960
136  TTCAGTGGCT  GGATCTCTGC  TCTGGTGGTG  CCGAAGGGCA  ACACAGCAGT  ATCCGTGCTC    1020
137  ATGCTGCTGG  TCGCCCTGCT  CTTCACTGGC  ATTGCTGTGC  TAGGAATTGT  CATGCTGAAA    1080
138  CGGATCCACT  CCTTATACCG  CCGCACAGGT  GCCAGCTTTC  AGAAGGCCCA  GCAAGAATTT    1140
139  GCTGCTGGTG  TCTTCTCCAA  CCCTGCGGTG  CGAACCAGC  CTGCCAATGC  AGCCGCTGGG    1200
140  GCTGCTGAAA  ATGCCTTCCG  GGCCCCGTGA  CCCCTGACTG  GGATGCCCTG  GCCCTGCTAC    1260
141  TTGAGGGAGC  TGAATTAGCT  CCCGTCCCTA  AGGTCTCTGG  GACTTGAGGA  GACATCACTA    1320
142  ACTGATGGCT  CCTCCGTAGT  GCTCCCAATC  CTATGGCCAT  GACTGCTGAA  CCTGACAGGC    1380
143  GTGTGGGGAG  TTACTGTGA  CTTAGTCCCT  CCATCAGGCC  ACACTGCTGC  CACCTCTCAC    1440
144  ACGCCCCAAC  CCAGCTTCCC  TCTGCTGTGC  CACGGCTGTT  GCTTCGGTTA  TTTAAATAAA    1500
145  AAGAAAGTGG  AACTGGAACT  G
146
147  (2) INFORMATION FOR SEQ ID NO: 3:
148
149  (i) SEQUENCE CHARACTERISTICS:
150      (A) LENGTH: 329 amino acids
151      (B) TYPE: amino acid
152      (C) STRANDEDNESS: single
153      (D) TOPOLOGY: linear
154
155  (vii) IMMEDIATE SOURCE:
156      (A) LIBRARY: BRSTNOT01
157      (B) CLONE: 412453
158
159  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
161  Met Ser Ala Phe Asp Thr Asn Pro Phe Ala Asp Pro Val Asp Val Asn
162  1          5          10          15
163  Pro Phe Gln Asp Pro Ser Val Thr Gln Leu Thr Asn Ala Pro Gln Gly
164  20          25          30
165  Gly Leu Ala Glu Phe Asn Pro Phe Ser Glu Thr Asn Ala Ala Thr Thr
166  35          40          45
167  Val Pro Val Thr Gln Leu Pro Gly Ser Ser Gln Pro Ala Val Leu Gln
168  50          55          60
169  Pro Ser Val Glu Pro Thr Gln Pro Thr Pro Gln Ala Val Val Ser Ala
170  65          70          75          80
171  Ala Gln Ala Gly Leu Leu Arg Gln Gln Glu Glu Leu Asp Arg Lys Ala

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,805

DATE: 06/08/2001

TIME: 11:46:26

Input Set : C:\PAOLA\09841805.txt

Output Set: N:\CRF3\06082001\I841805.raw

```

172          85          90          95
173 Ala Glu Leu Glu Arg Lys Glu Arg Glu Leu Gln Asn Thr Val Ala Asn
174          100          105          110
175 Leu His Val Arg Gln Asn Asn Trp Pro Pro Leu Pro Ser Trp Cys Pro
176          115          120          125
177 Val Lys Pro Cys Phe Tyr Gln Asp Phe Ser Thr Glu Ile Pro Ala Asp
178          130          135          140
179 Tyr Gln Arg Ile Cys Lys Met Leu Tyr Tyr Leu Trp Met Leu His Ser
180          145          150          155          160
181 Val Thr Leu Phe Leu Asn Leu Leu Ala Cys Leu Ala Trp Phe Ser Gly
182          165          170          175
183 Asn Ser Ser Lys Gly Val Asp Phe Gly Leu Ser Ile Leu Trp Phe Leu
184          180          185          190
185 Ile Phe Thr Pro Cys Ala Phe Leu Cys Trp Tyr Arg Pro Ile Tyr Lys
186          195          200          205
187 Ala Phe Arg Ser Asp Asn Ser Phe Ser Phe Phe Val Phe Phe Phe Val
188          210          215          220
189 Phe Phe Cys Gln Ile Gly Ile Tyr Ile Ile Gln Leu Val Gly Ile Pro
190          225          230          235          240
191 Gly Leu Gly Asp Ser Gly Trp Ile Ala Ala Leu Ser Thr Leu Asp Asn
192          245          250          255
193 His Ser Leu Ala Ile Ser Val Ile Met Met Val Val Ala Gly Phe Phe
194          260          265          270
195 Thr Leu Cys Ala Val Leu Ser Val Phe Leu Leu Gln Arg Val His Ser
196          275          280          285
197 Leu Tyr Arg Arg Thr Gly Ala Ser Phe Gln Gln Ala Gln Glu Glu Phe
198          290          295          300
199 Ser Gln Gly Ile Phe Ser Ser Arg Thr Phe His Arg Ala Ala Ser Ser
200          305          310          315          320
201 Ala Ala Gln Gly Ala Phe Gln Gly Asn
202          325

```

204 (2) INFORMATION FOR SEQ ID NO: 4:

206 (i) SEQUENCE CHARACTERISTICS:

207 (A) LENGTH: 2434 base pairs

208 (B) TYPE: nucleic acid

209 (C) STRANDEDNESS: single

210 (D) TOPOLOGY: linear

212 (vii) IMMEDIATE SOURCE:

213 (A) LIBRARY: BRSTNOT01

214 (B) CLONE: 412453

216 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

218 NCCGGAAGTG GAGGGTCTAC ACGAAGCGCC GCTGGGTCTG GGTGCCCCGA GGCAGCAGCG      60
219 TTCGCGGAGT TCGCCCGCTG GCCCCGATC ACCATGTCGG CTTTCGACAC CAACCCCTTC      120
220 GCGGACCCAG TGGATGTAAA CCCCTTCCAG GATCCCTCTG TGACCCAGCT GACCAACGCC      180
221 CCGCAGGGCG GCCTGGCGGA ATTCAACCCC TTCTCAGAGA CAAATGCAGC GACAACAGTT      240
222 CCTGTCACCC AACTCCCTGG GTCCTCACAG CCAGCGGTTT TCCAGCCATC AGTGGAACCA      300
223 ACCCAGCCGA CCCCCAGGC CGTGGTGTCT GCAGCCAGG CAGGCCTGCT CCGGCAGCAG      360
224 GAAGAACTGG ACAGGAAAGC TGCCGAGCTG GAACGCAAGG AGCGGGAGCT GCAGAACACT      420
225 GTAGCCAAC TGCATGTGAG ACAGAACAAC TGGCCCCCTC TGCCCTCGTG GTGCCCTGTG      480

```

RAW SEQUENCE LISTING

DATE: 06/08/2001

PATENT APPLICATION: US/09/841,805

TIME: 11:46:26

Input Set : C:\PAOLA\09841805.txt

Output Set: N:\CRF3\06082001\I841805.raw

```

226 AAGCCCTGCT TCTATCAGGA TTTCTCCACA GAGATCCCTG CCGACTACCA GCGGATATGC 540
227 AAGATGCTCT ACTATCTGTG GATGTTGCAT TCAGTGACTC TGTTTCTGAA CCTGCTTGCC 600
228 TGCCTGGCCT GGTTCCTGGG CAACAGCTCC AAGGGAGTGG ACTTTGGCCT CTCCATCCTG 660
229 TGGTTTCTGA TCTTCACTCC CTGTGCCTTC CTTTGTGTTG ACCGACCCAT CTATAAGGCC 720
230 TTTAGGTCCG ACAACTCTTT CAGCTTCTTT GTGTTCTTCT TTGTATTTT TTGTCAAATA 780
231 GGGATCTACA TCATCCAGTT GGTGGGCATC CCTGGCCTGG GGGACAGCGG TTGGATTGCA 840
232 GCCCTGTCTA CACTGGATAA TCATTCCCTG GCCATATCAG TCATCATGAT GGTGGTGGCT 900
233 GGCTTCTTCA CCTCTGTGTC CGTGCTCTCA GTCTTCCTCC TGCAGCGGGT GCACTCCCTC 960
234 TACCGACGGA CAGGGGCCAG CTTCCAGCAG GCCCAGGAGG AGTTTTCCCA GGGCATCTTC 1020
235 AGCAGCAGAA CCTTCCACAG AGCTGCTTCA TCTGCTGCCC AAGGAGCCTT CCAGGGGAAT 1080
236 TAGTCCTCCT CTCTTCTCTC CCCCTCAGCC TTTCTCTCGC CTGCCTTCTG AGCTGCACTT 1140
237 TCCGTGGGTG CCTTATGTGG TGGTGGTTGT GCCCAGCACA GACCTGGCAG GGTCTTGCC 1200
238 GTGGCTCTTC CTCTCCCTC AGCGACCAGC TCTCCCTGGA ACGGGAGGGA CAGGGAATTT 1260
239 TTTCCCCCTC TATGTACAAA AAAAAACAAA GCTCTCTTTC CTTCTCTGGT GATGGTTTGG 1320
240 TAGGATTCTT TTGTCTCTGG AAGCAGTGGG ACTGAAGTTC TCTTCGTCTT GTGCACACAC 1380
241 AGACACCCCC ACACAGTTGG GATCACAGGC TGACCTGGGC CCATCCCAGC TGGAGCTTTC 1440
242 TGCCAGGGTC CTGGGCCTTG ACTCCCCCAC CCTGCAGGCC TGGCCTGAAT CTGGCTTCTT 1500
243 AGACACAGCC CAGTCCTTCC TGCTGGGCT GGAATAAGC CTCTCACAGG TTCTGGTGGA 1560
244 CAGATCTGTT CCCCAGGTCA CTCCAGTGGT CTCCAGGCTT CCAGAGAAGG CTGGTTGCCT 1620
245 CAAGCTCTTC TCTGCCTCAT AAACGGATCC AGAGAAGGCT GGTTGCCTTA AGCTCTTCCC 1680
246 TGCTTCGTGT TCCTGAGAAA CGGATTAATA GCCCTTATC CCCCTGCACC CTCCTGCAGG 1740
247 GGATGGCACT TTGAGCCCTC TGGAGCCCTC CCCTTGCTGA GCCTTACTCT CTTCACTT 1800
248 TCTGAATGTA CAGTGCCGTT GGTGGGATT TGGGAGTGG AAGGGACCAA GGACACTGAC 1860
249 CCCAAGCTGT CCTGCCTAGC GTCCAGCGTC TTCTAGGAGG GTGGGGTCTG CCTGTCCTGG 1920
250 TGTGGTTGGT TTGGCCCTGT TTGCTGTGAC TACCCCCCCC CCTCCCCGAA CCGAGGGACG 1980
251 GCTGCCTTTG TCTTGCCCTC AGATGCCACC TGCCCCGCCC ATGCTCCCCA TCAGCAGCAT 2040
252 CCAGACTTTC AGGAAGGGCA GGACCAGCCA GTCCAGAACC GCATCCCTCA GCAGGGACTG 2100
253 ATAAGCCATC TCTCGGAGGG CCCCTAATA CCCAGTGGAG TCTGGTTCAC ACCCTGGGGG 2160
254 GTGTGTCACT GTGATGGGAC ACGTAGGAGT CCACCTTAA AACCAGCACC CTGTCCCTCG 2220
255 AGGCTGCCGA GTGGGTGTGT GGACTGGGGT GCCTTCCCAC AAAACTAGCC TCCGGCTCTG 2280
256 GGCCCGAGAC AGCCGCAGGC CCCAGCCACT GAATGATACT GGCAGCGGCT GGGGTTTTAT 2340
257 GAACTCCTTT CTGGTATTTT TTCCCTCTA TGTACAAATG TATATGTTAC GTCTCAATTT 2400
258 TTGTGCTTAA GTAAAAATAA AAACATTTTC AGAC 2434

```

260 (2) INFORMATION FOR SEQ ID NO: 5:

262 (i) SEQUENCE CHARACTERISTICS:

263 (A) LENGTH: 338 amino acids

264 (B) TYPE: amino acid

265 (C) STRANDEDNESS: single

266 (D) TOPOLOGY: linear

268 (vii) IMMEDIATE SOURCE:

269 (A) LIBRARY: GenBank

270 (B) CLONE: 487057

272 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

274 Met Ser Asp Phe Asp Ser Asn Pro Phe Ala Asp Pro Asp Leu Asn Asn
275 1 5 10 15
276 Pro Phe Lys Asp Pro Ser Val Thr Gln Val Thr Arg Asn Val Pro Pro
277 20 25 30
278 Gly Leu Asp Glu Tyr Asn Pro Phe Ser Asp Ser Arg Thr Pro Pro Pro
279 35 40 45

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/841,805

DATE: 06/08/2001

TIME: 11:46:27

Input Set : C:\PAOLA\09841805.txt

Output Set: N:\CRF3\06082001\I841805.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]